2002 Washington State Population Survey
Data Report

THE 2002 WASHINGTON STATE POPULATION SURVEY (SPS) is the third in a series of surveys designed to provide, biennially, a detailed profile of Washington State residents. The survey provides information on topics such as employment, income, education, immigration, health, and health insurance and borrows its structure and many of its questions from the national Current Population Survey (CPS). The original survey was constructed in 1998, and minor changes were made for the 2000 and 2002 surveys.

The SPS is a valuable complement to other reports and data resources addressing Washington's population. For example, although the federal Census was conducted in 2000, the SPS provides a consistent source of information at more frequent intervals than the decennial Census. Also, the March CPS measures income statewide on a year-to-year basis, but has a much smaller sample size and does not provide the regional income data found in the SPS.

A total of 6,842 households from the two separate samples completed the telephone interview in spring of 2002. The response rate was 47 percent for the general population sample and was 38 percent for the expanded sample. The average interview time was approximately 23 minutes. The interview questions were also translated into Spanish for the 2002 survey. OFM used population forecasts based on the 2000 Census figures for population by county as control totals to convert the raw statistical results of the survey into a recognizable and useful portrayal of the population of Washington State.

Survey Design

Sample Design

POPULATION

The population for this survey consists of all households located within the geographic boundaries of Washington State. Because the SPS was a telephone survey, only the households with telephones were potential subjects. The 2000 census shows that less than two percent of Washington households did not have telephones. Households on military bases and other group quarters - such as student dormitories, prisons, and nursing homes - were also excluded from this survey.[1] Since there is no universal list of all the households as defined above from which a random sample can be obtained, Gilmore used the random digit dialing (RDD) approach to obtain the required sample. The RDD approach is most commonly used to ensure equal probability of selection for each household with an activated telephone line, listed or not. Survey Sampling, Inc. (SSI) prepared the RDD sampling frame Gilmore used.

[1] See discussion of limitations.

SAMPLING

General Population Sample and Expanded Sample. Two separate samples were drawn for this survey. One was a random sample of all Washington State households, or the general population. The targeted number of completed interviews for the general population was 6,000. The second was an expanded sample of households in which the respondent was African American, Asian, Hispanic, or Native American. This expanded sample of minority groups enables data users to make inferences about characteristics of all major population groups. When examining the entire state population, responses from the expanded sample will be weighted to represent the incidence of these groups in the general population. The target for each of the minority groups identified above was 338 interviews from the general population sample and the expanded sample combined.

Since the RDD sample was inclusive of all state regions, it already provided a fair representation for each minority group. Our approach for expanding the sample of these populations was to identify the top five to ten census tracts for each minority group and to obtain a sufficient quantity of telephone numbers to ensure completion of the desired 338 completed interviews for each minority group.

Regional Stratification. The general population sample is stratified into eight geographic regions based on county of primary residence. The target completion for each region was 750. This regional grouping considered the similarities of economic and population characteristics among the 39 counties in Washington State. It was the result of consultation with legislative and other advisory groups for the State Population Survey.

Western Washington counties were grouped into five regions as follows:

Region 1: Island, San Juan, Skagit, Whatcom
Region 2: Clallam, Cowlitz, Grays Harbor, Jefferson, Klickitat, Lewis, Mason, Pacific, Skamania, Wahkiakum
Region 3: King
Region 4: Kitsap, Pierce, Snohomish, Thurston
Region 5: Clark

Eastern Washington counties were grouped into three regions as follows:

Region 6: Adams, Asotin, Chelan, Columbia, Douglas, Ferry, Garfield, Grant, Kittitas, Lincoln, Okanogan, Pend Oreille, Stevens, Walla Walla, Whitman
Region 7: Spokane
Region 8: Benton, Franklin, Yakima

Questionnaire Design

The initial draft of the 1998 questionnaire was based on the March CPS questionnaire. In addition to the CPS questionnaire, the OFM State Population Survey group solicited additional questions for key subject areas. The draft was reviewed by a group of more than eighty individuals representing different organizations. Their comments were collected and reviewed by OFM, and many of their

recommendations were incorporated into the final questionnaire. In 2002, OFM staff reviewed the 2000 questionnaire and made minor changes clarifying medical insurance questions.

II. Survey Administration

The survey was administered by the Gilmore Research Group. Before the full-scale fielding, a pretest of 107 cases was conducted. The full-scale fielding started on February 5, 2002, and the interview phase was completed on June 4, 2002. The average interview time was approximately 23 minutes.

Advance Letter. To obtain full cooperation from the potential respondents, Gilmore sent an advance letter to about 7,900 households to announce the survey and explain its purpose. Matching phone numbers with existing directories generated the addresses of these households.

Response Rates. A total of 6,842 households completed the survey. Response rates were calculated separately for the general population sample and the expanded sample. The Council of American Survey Research Organizations recommends a calculation method that involves a total account of the sample dispositions and an estimation of eligible households from non-contact cases. According to this calculation method, the response rate for the general population sample is 47 percent and for the expanded sample is 38 percent. A separate technical report discusses in detail the sample disposition and calculation of the response rates for this survey.

III. The Analysis Data File

The analysis data file consists of 235 variables that were either extracted from the original survey data file or constructed at OFM. In the analysis file, the data are arranged so that each person's data are on an individual record. Thus, a household with five members has five records. The file contains 17,437 persons from 6,842 households.

The analysis data file is available in both SAS format and Excel format. It can be downloaded from the OFM Web-page for SPS. The URL address is http://www.ofm.wa.gov/sps/index.htm.

IV. Data Tabulations

The data tabulations are presented in two forms, one for categorical and one for continuous variables. The former uses frequency analysis and the latter means analysis.

Each record is weighted and the eight regional values are presented alongside the state value. Also, all tabulations include the variable name, variable label, and the universe for that variable. A subtitle indicates whether the variable is a person variable, a household or respondent-only variable, or a family variable.

A variable with a frequency analysis runs through at least two pages (indicated as Part 1 of 2 and Part 2 of 2). The first page lists the state value and the numbers for Regions 1 through 4. The second page lists the state total again and the numbers for Regions 5 through 8. The number of pages will increase by an increment of two depending on the number of data levels in a variable. Under the state and region headings, the weighted frequency counts and percentages for each data level are listed. Other information in the frequency tables includes a maximum margin of error at the 95 percent confidence level for the state and for each of the eight regions.

Margin of Error. Caution should be used in interpreting tabulations that contain small values with a relatively large margin of error. Take for example the question: *In which county did you [the respondent] live one year ago, if not in current Washington county?* The weighted tabulation shows that a state total of 1,992 people lived in Chelan County one year ago. They constituted about 1.3 percent of individuals who were reported to have moved between counties in Washington within the past year. However, the ±1.7 percent margin of error indicates that we are only reasonably confident that the true number of former Chelan County residents is somewhere between zero and approximately 3.0 percent. A common practice to reduce the standard error in such situations is to combine the data levels with fewer categories. In this particular example, instead of individual states, regions can be created.

For each of the means analyses, the numbers for the eight regions and the state total are all listed on one page. This type of table includes the following statistics:

- Total non-missing observations
- Mean
- Minimum
- Maximum
- Median
- Total observations
- Total missing observations
- Sum of weights
- Lower limit of 95 percent confidence interval
- Upper limit of 95 percent confidence interval

It should be pointed out that because of extremely high values in some of the continuous variables, the mean tends to be skewed. In such cases, the median is a better measure of the central tendency.

V. Limitations

Since this survey was a telephone survey, households without telephones were excluded. This non-coverage is, however, quite small. Statewide, the percent of households without telephones was less

than 1.4 percent according to the 2000 census. While there exists the risk of systematically missing some people in a telephone survey, most researchers do not consider it to be a serious problem.[2]

Another limitation common to all surveys is "non-responses." This term refers to households that refuse to participate in the survey. The response rate in this survey is 47 percent for the general population sample and 37 percent for the expanded sample. As in all surveys, there is a potential distortion in the results if the characteristics of the non-responding households are systematically different than those of the responding households. A common practice to partially compensate for the non-response error is to post-stratify the survey based on known population characteristics,[3] which was done in this project.

An examination of the responses suggests that the degree of distortion due to non-responses is small. OFM examined frequencies, means, and medians of selected key variables in the data set and compared the results with alternative data sources. For example, wage data from the survey was compared with wage information from the state Unemployment Insurance System. In virtually all cases where survey data were compared with alternative data sources, the results were very similar. The issue of non-response and comparisons between survey results and alternative data sources for key variables will be discussed in a forthcoming technical report.

A third limitation in this survey is the difference between the design and the post-stratification with respect to group-quarters populations. While the design called for exclusion of group-quarters populations, in the post-stratification process, the group-quarters population could not be separated from the general population estimates. Thus, the survey data were weighted to the entire state population.

---

[2] Folz, D.H. (1996). *Survey Research for Public Administration*. Thousand Oaks, CA: Sage.
Frey, J.H. (1989). *Survey Research by Telephone*. Newbury Park, CA: Sage.

[3] Lavrakas, P.J. (1993). *Telephone Survey Methods: Sampling, Selection, and Supervision*. Newbury Park, CA: Sage.